A Requirement-driven Approach for Designing Data Warehouses

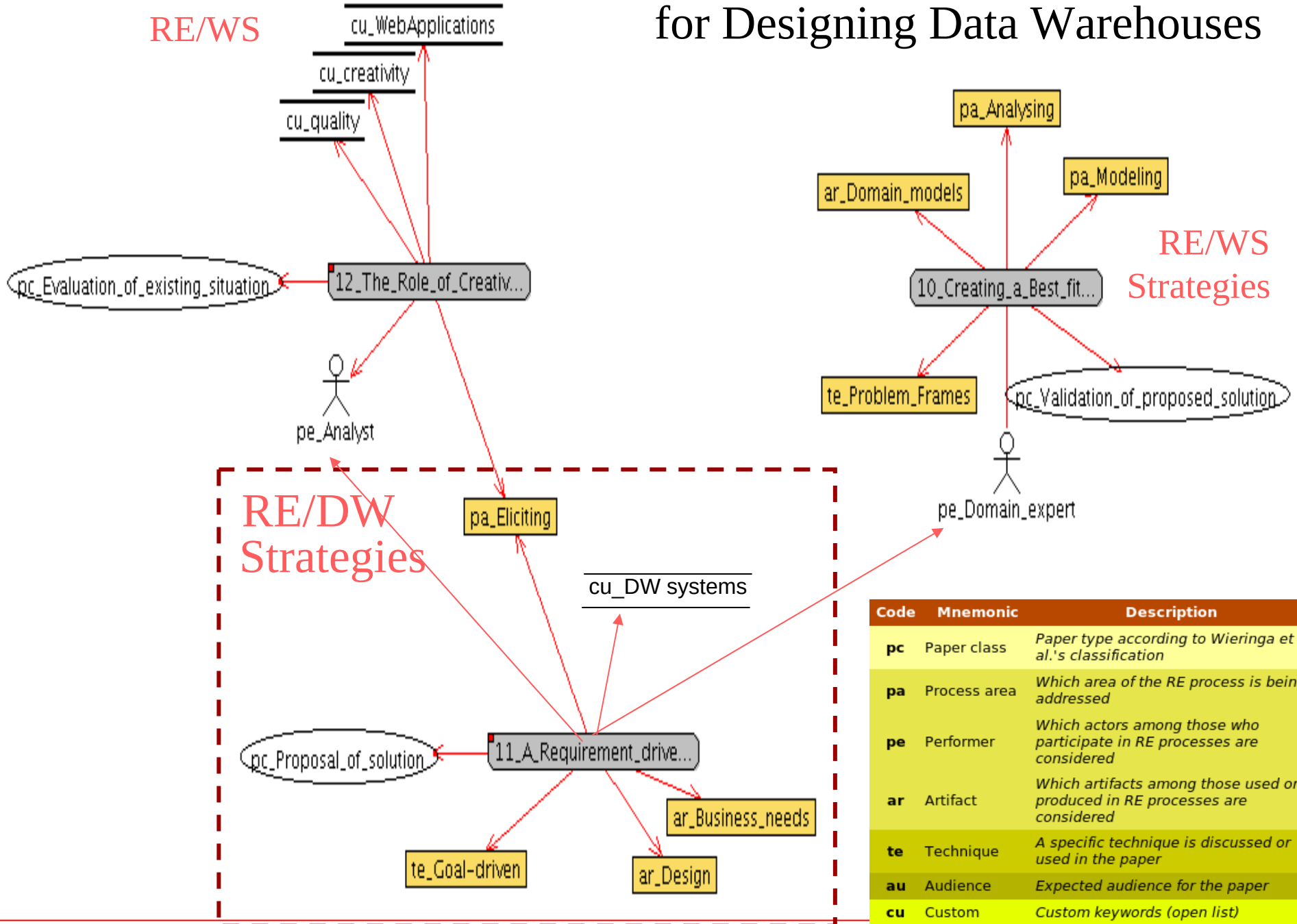
- CADWA -

Inès GAM *Dr. Camille SALINESI*

Centre de Recherche en Informatique

Université Paris1 – Panthéon Sorbonne

A Requirement-driven Approach for Designing Data Warehouses



Code	Mnemonic	Description
pc	Paper class	*Paper type according to Wieringa et al.'s classification*
pa	Process area	*Which area of the RE process is being addressed*
pe	Performer	*Which actors among those who participate in RE processes are considered*
ar	Artifact	*Which artifacts among those used or produced in RE processes are considered*
te	Technique	*A specific technique is discussed or used in the paper*
au	Audience	*Expected audience for the paper*
cu	Custom	*Custom keywords (open list)*

Context and Motivation

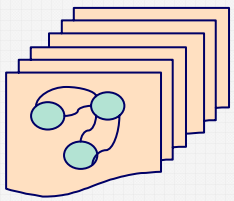
	Transactional System	*Decisional System*
Users	Operational agents	Decision maker
	Thousands	Hundreds
Design	Operationally oriented	Subject oriented
	Relatively static	Change
Data	For a « certain » period	historized
	Detailed	Aggregated
	Updated	Recalculated
Treatments	Simple requests	Complex requests
	Repetitive	Varying
Accessed tuples	dozens/hundreds	Thousands
Data Bases	MB/GB	GB/TB

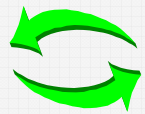
Evaluation

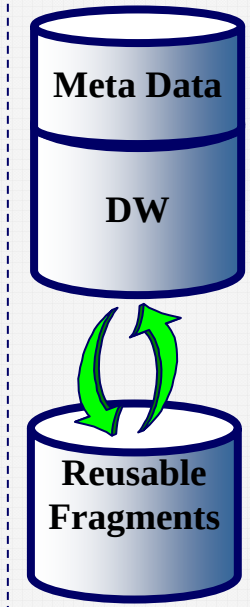
CADWA approach

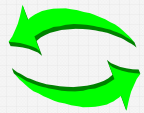
Strategies

Decisions









Informations

High level requirements

Schiefer, Böhnlein, Paim, winter, Prakash, Mazon

Overview of CADWA

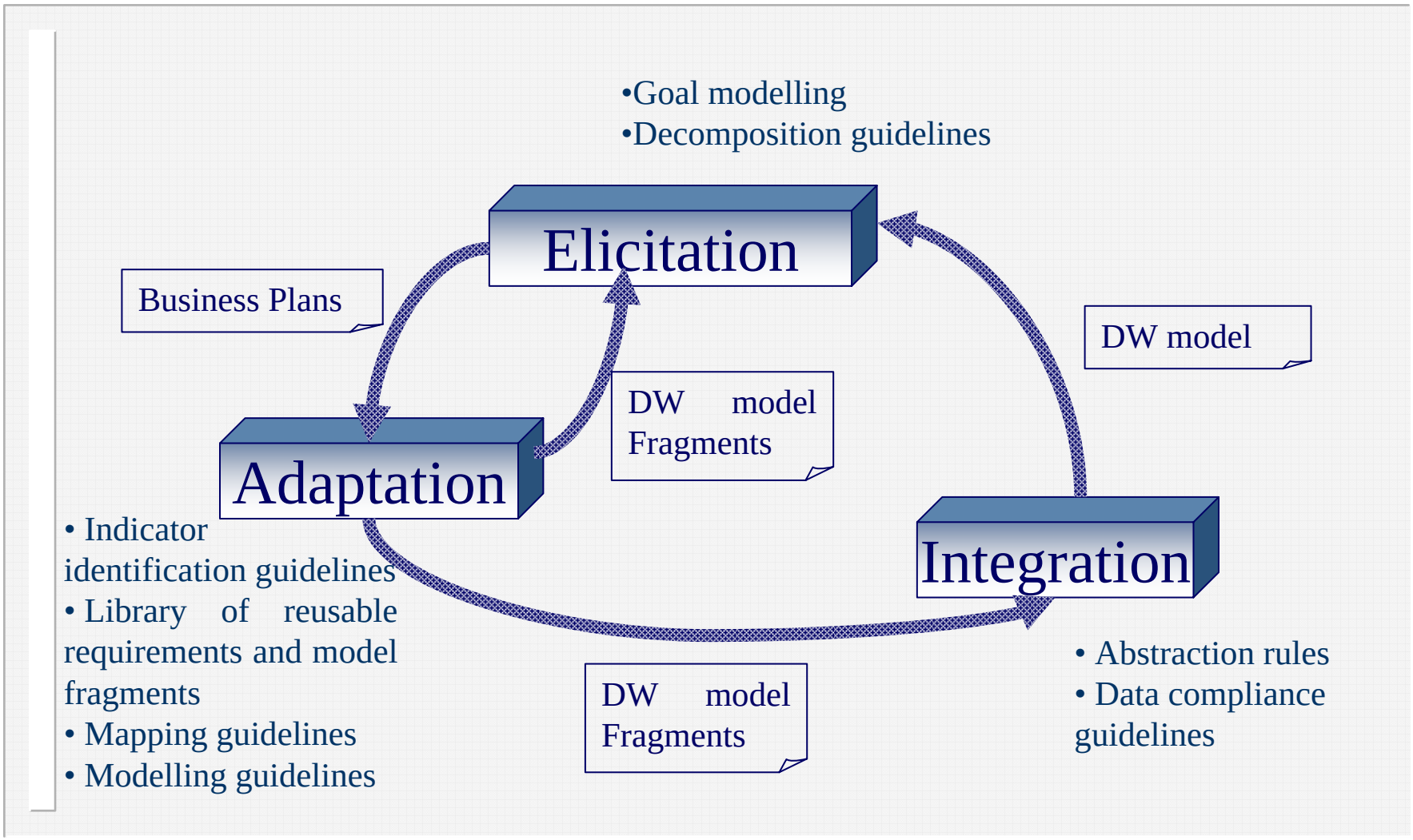


Illustration Case-Study presentation

- Adjust price position, especially in the hypermarkets in France.
- Improve growth and profitability of international business.
- Rationalize business portfolio
- Strengthen financial position

Hypermarkets | Supermarkets | Hard discount | Convenience and other businesses

Overview of CADWA

- Requirement Elicitation Phase -

Elicit Requirements

Input:

-Informal DW user requirements

Output:

-Organization Business plan (OBP)

-Decision-maker macro business plan (MaBP)

-Decision-maker micro business plan (MiBP)

-Action Plan (AP)

Sub activities:

-Elicitate objectives

-Define OBP

-Distribute OBP among users

-Define MiBP

-Operationalize actions

Elicitate Requirements guidelines

Adaptation

adapting

Integration

System requirements

OBP OBP OBP OBP

Define OBP

MaBP MaBP MaBP

Distribute OBP among users

MiBP MiBP MiBP

Define MiBP

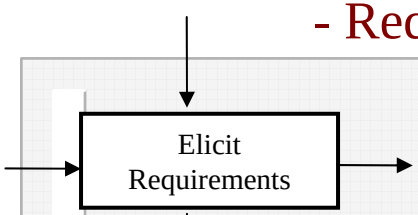
AP AP AP

Operationalize actions

Plans according to high level requirements

Overview of CADWA

- Requirements Elicitation phase : Design of business plans -



Input:

-Informal DW user requirements

Output:

-Organization Business plan (OBP)

-Decision-maker macro business plan (MaBP)

-Decision-maker micro business plan (MiBP)

-Action Plan (AP)

Sub activities:

-Elicitate objectives

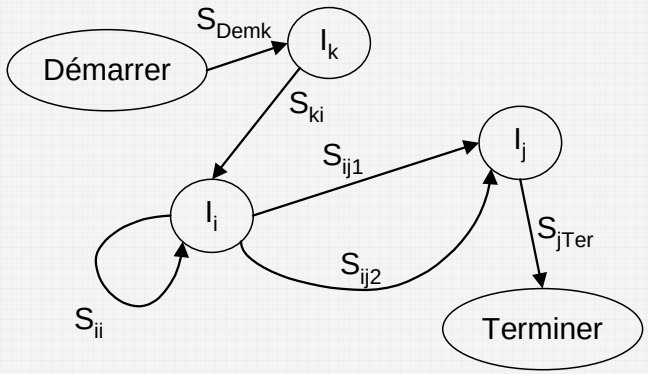
-Define OBP

-Distribute OBP among users

-Define MiBP

-Operationalize actions

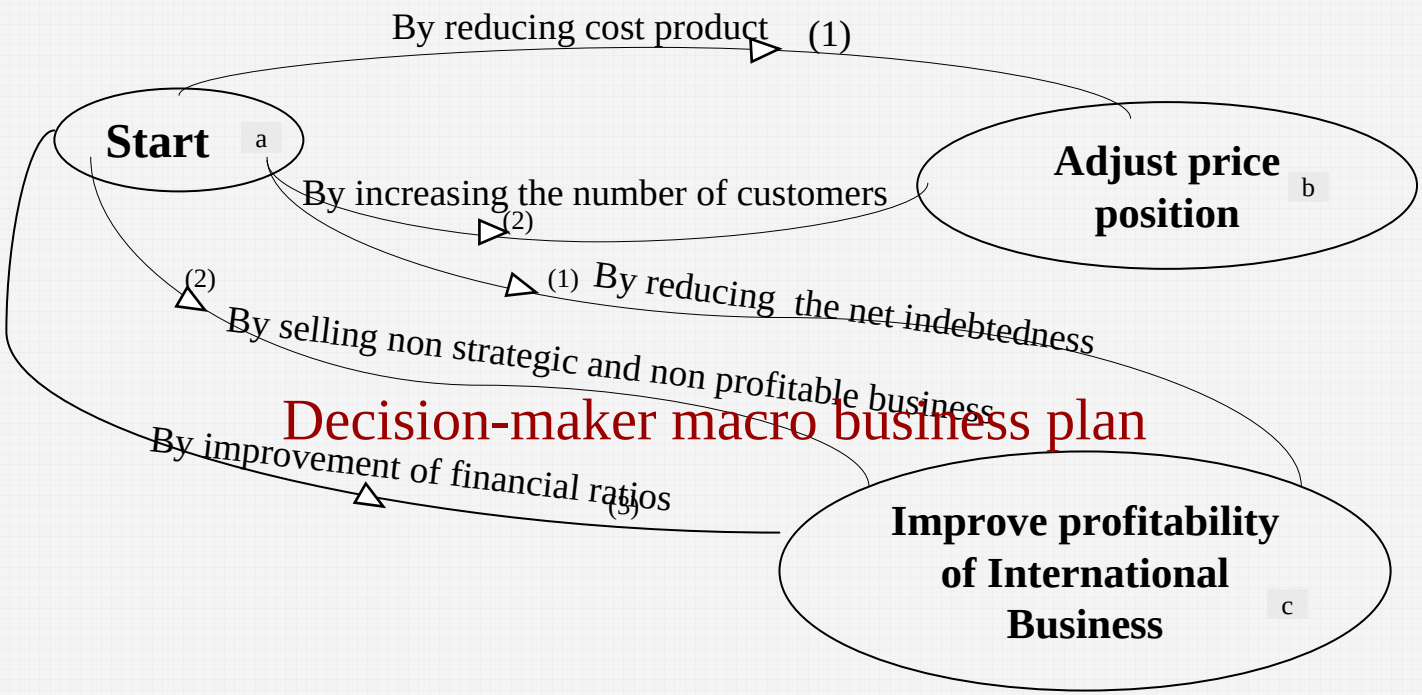
Elicitate Requirements guidelines



S0: Demarrer,I_k, S_{Demk}
S1: I_i,I_j,S_{ij1}
S2: I_i,I_j, S_{ij2}
S3: I_i, I_i, S_{ii}
S4: I_k,I_i,S_{ki}
S5: I_j,Terminer,S_{jTer}

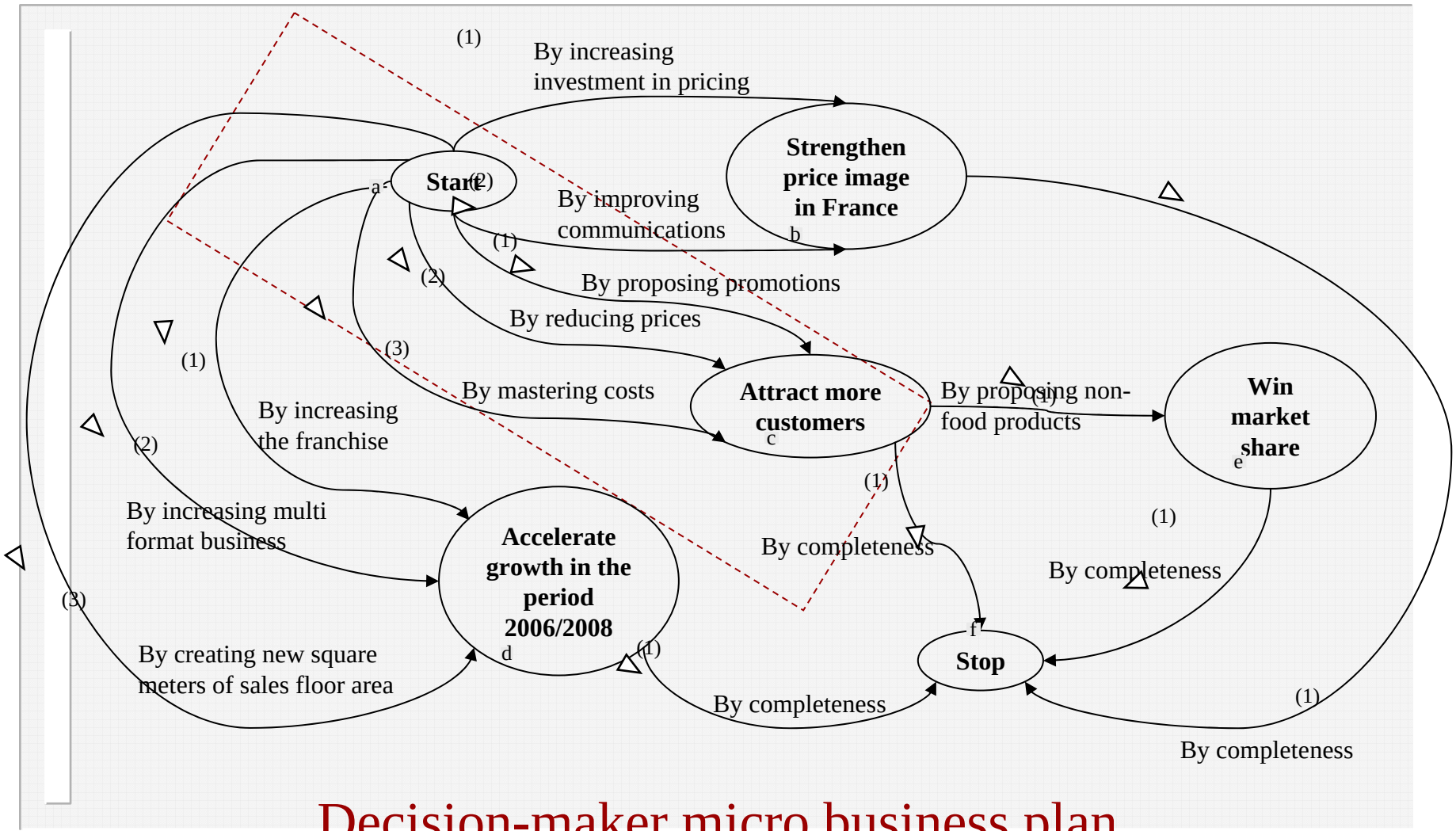
– A map is a labelled directed graph
– An *intention* (goals to achieve or maintain) as nodes
– A strategy is an edge
–A *section* is a triplet <*Ii*, *Ij*, *Sij*> where *Ii* is the source intention, *Ij* the target intention and *Sij* the strategy to attain when *Ii* has been achieved.

Illustration



Organization Business Plans

Illustration



Decision-maker micro business plan

Illustration

–**AI1**: Sell $_{verb}$ 40% more $_{quantity}$ childcare products $_{object}$ in French hypermarkets $_{location}$ during the Christmas period $_{time}$.

–**AI2**: Reduce $_{verb}$ products returns $_{object}$ in hypermarkets $_{location}$ after the promotion periods $_{time}$.

–**AI3**: Reduce $_{verb}$ products returns $_{object}$ from market segment 4 $_{source}$ to stores $_{destination}$ on after the promotion periods $_{time}$.

–**AI4**: Sell 042-0000073465-3 $_{object}$ in co with promotional conditions

–**AI5**: Push $_{verb}$ the gross sales $_{result}$ of baby wipes $_{object}$ with promotions on <wipe+diaper> packages $_{means}$.

–**AI6**: Propose $_{verb}$ special sales $_{result}$ to families $_{beneficiary}$.

–Define effective action intention plan

–Define action intention plan according to:

–Verb <Target> [<Parameter>]*

Action intention Plan

Overview of CADWA

- Design Phase -

Design DW Fragment Model

Input:

-Reusable DM model fragments

Output:

-DW model fragment (DwMf)

Sub activities:

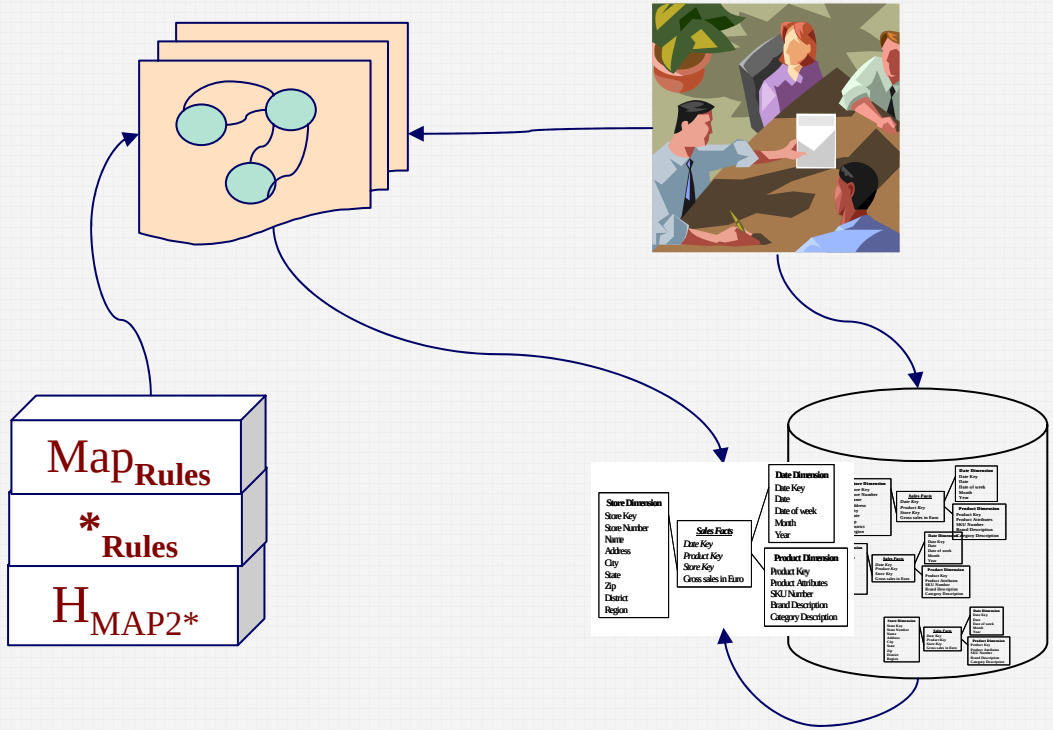
-*Extract modeling indications*

-*Select reusable DM*

-*Propose* DwMf

-*Adapt DM*

Design DW Fragment model guidelines



DW Fragment according to high level requirements

Illustration

SR1: In the dimensional models, the facts tables express the relation of one to many between the dimension tables
SR2: A line in a fact table corresponds to several me
measure is an attribute in a fact table. All the measures
fact table
SR3: A fa
only for ce
most usefu
SR4: Atal
SR5: Attr
analyses ir

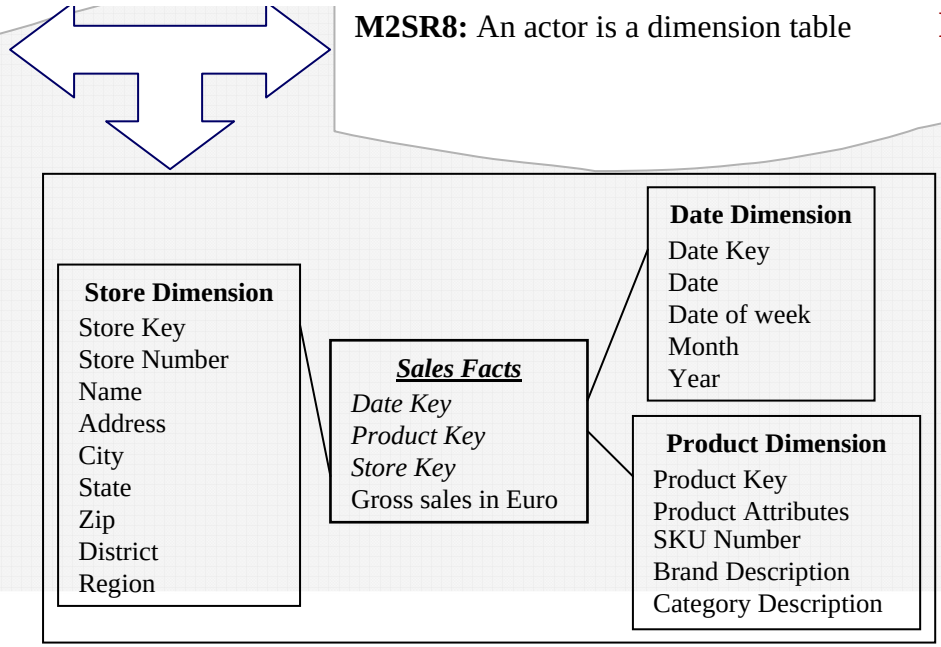
M2SR1: The parameter "destination" of direction is a dimension table or dimension
attribute

Sell (verb) 40% more (quantity) childcare products (object) in French hypermarkets (location) during the Christmas period (time).

Sell (verb) a quantity (quantity) of products (object) in hypermarkets (location) during a certain period (time).

M2SR8: An actor is a dimension table

Map to Star Rules



DW Fragment model

Overview of CADWA

- Integration Phase -

Integrate Models

Input:

-DW model fragment

-Legacy DW models

-Existing DB models

Output:

-New DW model

Sub activities:

- *Abstract legacy DW*

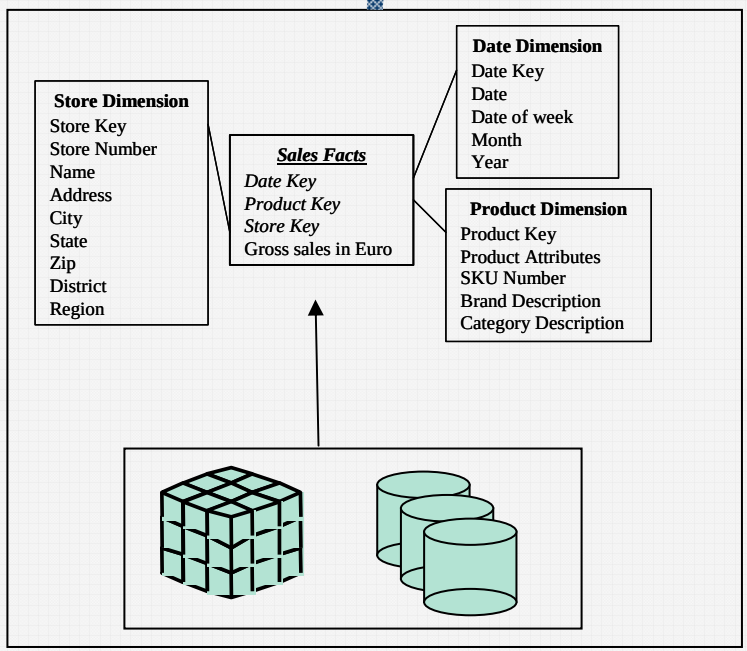
-*Abstract DB*

-*Generate new DW model*

Integrate models guidelines

Elicitation

System requirements



Generate new DW model

Abstract DB and legacy DW

DW model according to requirements

Illustration

Dimension Dé mographique Dé partement
Clé_dé mographique_dé partement
Population_totale
Population_moins_de_5ans
%population_moinsde_5ans
Population_moins_de_18ans
%Population_moins_de_18ans
Population_de_65ans et_plus
% Population_de_65ans_et_plus
Population_fé minine
% population_fé minine
Population_masculine
% population_masculine
Nombre_de_bacheliers
Nombre_de_diplômé s_é tudes_supé rieures
Nombre_de_logement
% prprié taires_de_leur_logement..

Dimension Client
Clé _Client
Code_Client
Appelation_client
Pré nom_client
Nom_Client
Ville_Client
Dé partement_Client
Clé_dé mographique_dé partement
Ré gion_client…

Dimension Produit
Clé _produit
Description_du_ produit
Numé ro_unité _de_stock_produit
Description_marque_produit
Description_caté gorie_produit
Description_de_sous_caté gorie
Description_rayon
Description_type_emballage
Taille_emballage
Poids_produit
Type_stockage_produit
Type_de_duré e_sur_étagère
Largeur_sur_étagère
Hauteur_sur_étagère
Profondeur_sur_étagère…

Dimension Promotion
Clé _promotion
Nom_promotion
Type_de_ré duction_de_prix
Type_de_mé dia_de_l' annonce
Type_d' anonce
Tye_de_pré sentation
Type_de_coupon
Nom_du_mé dia_de_l' anonce
Fournisseur_de_la_pré sentation
Coû t_de_la_promotion
Date_dé ut_promotion
Date_fin_promotion..

Faits Inventaire des ventes
Clé _client
Clé _magasin
Clé _date
Clé _produit
Unité s_vendues
Prix
Tax

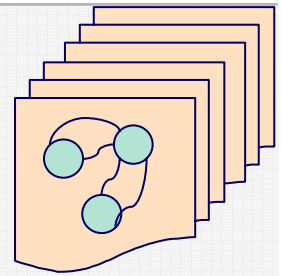
Faits retours
Clé _Client
Clé _magasin
Clé _date
Clé _produit
Clé _retour_trx
Clé _origine_trx
Code_raison
Unité s_retournées
Prix
Taxe_ remboursement
Montant_remboursement

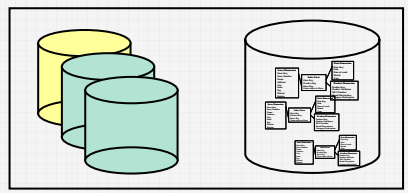
Faits vente journalière
Clé _Date
Clé _produit
Clé _magasin
Ventes (euro)..

Faits de transactions terminal point devente
Clé _Date
Clé _magasin
Clé _prmotion
Clé _produit
Numé ro_de_transaction_terminal point_vente
Quantité _vendue
Ventes (euro)
Coûts (euro)
Marge_brute (euro)

Dimension Date
Clé _Date
Date
Date_description complète
Jour_de_la_semaine
Numé ro_de_jour_ds_l'é poque
Numé ro_de_semaine_ds_l'é poque
N° _jour_ds_l'anné e_calendrier
N° _jour_ds_le_mois_calendrier
N° _jour_ds_le_mois_exercice
N° _jour_ds_l'anné e_exercice
Indicateur_dernier_jour_de_la_semaine
Indicateur_dernier_jour_du_mois
Date_de_la_fin_de_a_semaine_calendrier
N° _sem.ds_l'anné e_calendrier
Nom_du_mois_du_calendrier
N° _de_mois_de_l'anné e_calendrier
Anné e_mois_cal._(AAAA MM)
Trimestre_calendrier
Anné e_trimestre_calendrier
Semestre_calendrier
Anné e_calendaire
Semaine_exercice
N° _semaine_ds_l' exercice
Mois_de_l' exercice_
N° _mois_ds_l' exercice
Anné e_mois_exercice
Trimestre_de_l' exercice
Anné e_Trimestre_exercice
Semestre_de_l' exercice
Indicateur_de_jour_férier
Indicateur_de_jour_de_la_semaine
Saison_de_vente
Evé nement_majeur
Dat_SQL…

Dimension Magasin
Clé _magasin
Nom_magasin
N° _adresse_magasin
Rue_adresse_du_magasin
Code_postal_adresse_du_magasin
Zone_commerciale_du_magasin
Ré gion_commerciale_du_magasin
Manager_magqsin
Type_service_financier
Superficie_vente_magasin
Date_premiè re_ouverture
Date_dernière_ré novation...





Conclusion

- Overview of *CADWA* method

- Formalizing the *CADWA* method
- Development of a support tool
- Evaluation by developing a questionnaire on DW development practice

–Which quality features are addressed by the paper?

Proposing CADWA method for designing DW by analysing decision makers high level requirement.

–What is the main novelty/contribution of the paper?

Analysing high level requirement of decision makers in order to integrate them while designing DW

How will this novelty/contribution improve RE practice or RE research?

Process decision models are hardly modelized because of the "confidentiality" of decision makers. So usage world of DW can't be considered as in traditional RE. Our proposal is deal with usage world by focusing in subject world.

–What are the main problems with the novelty/contribution and/or with the paper?

Formalisation is needed and validation according to complex real life problems.

–Can the proposed approach be expected to scale to real-life problems?

Yes. Some initial evaluation are made and we expect to make more intensive evaluation in order to confrontate our method to more complex real life problems